DOCUMENT RESUME

ED 413 099 PS 025 966

AUTHOR Gamse, Beth C.; Conger, Dylan; Elson, Dean; McCarthy, Maria
TITLE Follow-Up Study of Families in the Even Start In-Depth Study. Final Report.
INSTITUTION Abt Associates, Inc., Cambridge, MA.
SPONS AGENCY Department of Education, Washington, DC. Planning and Evaluation Service.
PUB DATE 1997-00-00
NOTE 39p.
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Disadvantaged Environment; *Disadvantaged Youth; Early Childhood Education; *Family Literacy; Intervention; *Literacy Education; Low Income Groups; Outcomes of Education; Program Effectiveness; Program Evaluation; School Readiness; Young Children
IDENTIFIERS *Even Start

ABSTRACT

The federal Even Start program was implemented to improve the educational opportunities of low-income children and adults by integrating early childhood education, adult education, and parent education into a unified family literacy initiative. This study was designed to assess the impact of Even Start on children's early school performance several years after their participation in the program and, secondarily, to characterize the schools in which the Even Start and comparison children were enrolled. Data were collected on a subset of children from the In-Depth Study of the first national evaluation of Even Start. Of the 179 children included in that study, 128 (72%) were involved in the Follow-Up study. The majority of the students were in the first and second grades during the 1994-95 school year. The medium-term effects of Even Start participation were assessed using the following measures: school grades; achievement test data; school attendance; special education placement; Title I placement; participation in transitional classrooms, summer school programs, and bilingual education; and grade retention history. The study found little difference between former Even Start participants and a control group except that Even Start participants were less likely to be tardy arriving at school. Given that previous research has demonstrated the positive long-term effects of early childhood education programs, these results might be explained by the absence of the following features in this study: (1) a longer interval between the treatment and subsequent follow-up; (2) the use of comparable grade data or achievement test scores; and (3) the capacity to collect new data from children as well as from teachers or parents. While it is possible that the Even Start program does not have medium- or long-term effects, it is also possible that these features needed to be included in the study design to detect meaningful differences. (Contains 24 references and 3 appendices which detail methodology and family characteristics.) (LPP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 413 099

Abt Associates Inc.

55 Wheeler Street
Cambridge, Massachusetts
02138-1168

617 492-7100 *telephone*
617 492-5219 *facsimile*

Hampden Square, Suite 500
4800 Montgomery Lane
Bethesda, Maryland
20814-5341

301 913-0500 *telephone*
301 652-3618 *facsimile*

101 North Wacker Drive
Suite 400
Chicago, Illinois
60606-7301

312 332-3300 *telephone*
312 621-3840 *facsimile*

Follow-Up Study of Families in the Even Start In-Depth Study

Final Report

1997

Prepared by

Abt Associates Inc.
Cambridge, MA

Prepared for

U.S. Department of Education
Planning and Evaluation Service

PS 025966

BEST COPY AVAILABLE

FOLLOW-UP STUDY OF FAMILIES IN THE EVEN START IN-DEPTH STUDY

Abstract

The Follow-Up Study of Families in the Even Start In-Depth Study was designed to assess the impact of Even Start on children's early school performance several years after their participation in the program. The study found little difference between former Even Start participants and a control group with the exception that Even Start participants were less likely to be tardy from school. Findings from an earlier evaluation of the Even Start program indicated that there were short-term gains for children on some measures of school readiness. Related research on other early childhood education programs also suggests that there are positive effects of participation on school-related measures. This study assessed the effects of Even Start on children in grades 1 and 2 about three years after participating, using a variety of measures. It was produced by Abt Associates, Inc., under contract to the U.S. Department of Education's Planning and Evaluation Service (PES).

BEST COPY AVAILABLE

FOLLOW-UP STUDY OF FAMILIES IN THE EVEN START IN-DEPTH STUDY

FINAL REPORT

Beth C. Gamse
Dylan Conger
Dean Elson
Maria McCarthy

1997

Tracy Rimdzius
Project Monitor

TABLE OF CONTENTS

		Page
Executive Summary		v
Chapter One	**Introduction**	1
	Even Start Family Literacy Program Background	1
	Even Start Program Design	2
	Summary of the First National Even Start Evaluation Design	2
	Highlights of Findings from the In-Depth Study of the First Evaluation	3
	Research Questions for the Follow-Up Study	6
	Design of the Follow-Up Study	6
Chapter Two	**Hypothesized Effects of Even Start**	8
	Summary	11
Chapter Three	**Findings from the Follow-up Study**	12
	School Environments	12
	Student Demographics	12
	Participation in Special Programs/Services	12
	Student Attendance and Performance	15
	Achievement Tests	15
	Summary	15
References		18
Methodological Appendices		
Appendix A:	**Finding Families**	A-1
	Preparation of Family Fact Sheets	A-1
	Initial Contact with Sites/Local Recruiters	A-1
	Locating Families and Obtaining Consent	A-2
	Data Collection	A-2
Appendix B:	**Measurement and Analysis**	A-4
	Measures Used in the Follow-Up Study	A-4
	Analytic Approach to Assessing Program Effects	A-6
Appendix C:	**Description of Even Start Follow-Up Study Children and Families at Pretest**	A-9
	Demographic Characteristics of In-Depth Study Families at Pretest	A-10
	Literacy Level of Adults at Pretest	A-11
	Demographic Characteristics of Children at Pretest	A-12
	School Readiness and Literacy of Children at Pretest	A-13

LIST OF EXHIBITS

		Page
Exhibit 1.1	Even Start Conceptual Model	4
Exhibit 1.2	Number of Even Start and Comparison Children by Project Site	7
Exhibit 3.1	Percent of Schools with Title I Services 1994-1995 School Year	13
Exhibit 3.2	Children Participating in Title I Services 1994-1995 School Year	14
Exhibit 3.3	Children Participating in Available Special Programs 1994-1995 School Year	14
Exhibit 3.4	Attendance and Tardiness Rates	16
Exhibit 3.5	Average Grades in Reading, English, and Math	16
Exhibit C.1	Demographic Characteristics of Families at Pretest	A-10
Exhibit C.2	Pretest Scores on Comprehensive Adult Student Assessment System (CASAS)	A-11
Exhibit C.3	Children's Demographic Characteristics at Pretest	A-12
Exhibit C.4	Pretest Scores on Preschool Inventory (PSI) and Peabody Picture Vocabulary Test-Revised (PPVT-R)	A-13

EXECUTIVE SUMMARY

The *Follow-Up Study of Families in the Even Start In-Depth Study* was designed to assess the impact of Even Start on children's early school performance several years after children participated in Even Start. Findings from an earlier evaluation of the Even Start program indicate that there are short-term gains for children on some measures of school readiness, and related research on other early childhood educational programs also suggests that participation has positive effects on school-related measures. Because Even Start is hypothesized to improve children's school performance, federal, state, and other policy makers are especially interested in learning about the effects of the program once children have entered elementary school. The purpose of this follow-up study is to address the following questions:

- What are the effects of the program on children two to three years after program participation?
- How are Even Start children performing in school, now that they are in the primary grades?

The Follow-Up Study builds upon the strength of the In-Depth Study (IDS) design; in five IDS sites, families were randomly assigned either to Even Start or a comparison group. The current study focuses exclusively on those children whose families were participants in the random assignment component of the earlier evaluation. In order to assess children's school performance, we examined children's records on attendance, grades, scores on standardized achievement tests, grade retention, participation in special programs such as Title I or other compensatory education programs, and placement in special education services. Our analyses are based upon data from the 1994-95 school year, which represented the last completed school year at the time of data collection in the spring of 1996.

Our analyses indicate that there are few, if any, differences between Even Start and comparison children on the available indicators. The only significant difference is on tardy arrival to school: Even Start children are significantly less likely to arrive late than are comparison children. The current lack of findings is, however, consistent with research findings in many other rigorously designed studies of preschool educational programs. In many cases, differences between program and comparison children were observed much later in children's school careers and were observed using a wider range of measures than were available during the Follow-Up Study.

Chapter One
INTRODUCTION

The first section of this chapter briefly describes the Even Start Family Literacy Program. It is followed by a description of the mandate for the first National Even Start Evaluation and the short-term findings from that evaluation. The last section of this chapter lists the research questions and describes the design of the Even Start Follow-Up Study.

Even Start Family Literacy Program Background

The Even Start Family Literacy Program was originally authorized by the Elementary and Secondary Education Act of 1965 (ESEA) as amended by the Hawkins-Stafford Elementary and Secondary School Improvement Amendments of 1988, Part B of Chapter 1 of Title I (P.L. 100-297). The Even Start program was reauthorized in 1994 by the Improving America's Schools Act (P.L. 103-382), as part B of Title I of the ESEA. This description of Even Start refers to the reauthorized law. Projects were not required to implement changes made by that law, however, until Program year 1995-96. According to the 1994 legislation, the Even Start program is intended to:

> "...help break the cycle of poverty and illiteracy by improving the educational opportunities of the Nation's low-income families by integrating early childhood education, adult literacy or adult basic education, and parenting education into a unified family literacy program...The program shall (1) be implemented through cooperative projects that build on existing community resources to create a new range of services; (2) promote achievement of the National Education Goals; and (3) assist children and adults from low-income families to achieve to challenging State content standards and challenging State student performance standards."
> (P.L. 103-382, Sec. 1201).

To be eligible for Even Start under the reauthorized law, a family must have (a) an adult who is eligible for adult education programs under the Adult Education Act, or is within the state's compulsory school attendance age, and (b) have a child less than eight years of age. Even Start projects are required to provide participating families with an integrated program of early childhood education, adult literacy and basic education, and parenting education. The program's design is based on the notion that these components build on each other and that families need to receive all three services, not just one or two, in order to effect lasting change and improve children's school success. As a "family-focused" rather than parent- or child-focused program, Even Start has two interrelated goals:

- to help parents become full partners in the education of their children; and
- to assist children in reaching their full potential as learners.

To achieve these goals, Even Start began as a demonstration program administered by the U.S. Department of Education (ED) that provided school districts with four-year discretionary grants for family literacy projects in 1989. In 1992, the program, while remaining a competitive discretionary grant program, became primarily administered by the states, although two small set-asides remain for direct federal grants for Migrant Education projects and grants to Indian tribes and tribal organizations. In addition, the reauthorized law authorizes one grant in a prison that houses women and children, and grants for statewide family literacy initiatives.

According to the Even Start statute, when the program is funded for $50 million or more per year, it must be administered at the state level. Each state's share of Even Start funds is based on its proportion of funds under the Title I Part A LEA Program. States hold grant competitions and make subgrant awards. The statute specifies that each Even Start subgrantee must receive a minimum of $75,000 per year.

Even Start Program Design

The Even Start legislation contains language setting forth the major elements that must be the basis of each Even Start local project. The legislation allows grantees flexibility in devising projects to meet local needs but all projects are required to offer three core services:

- **Adult education and adult literacy**: high-quality instructional programs[1] for adults to promote adult literacy [including adult basic education (ABE), adult secondary education (ASE), English-as-a-second language (ESL), and preparation for the General Education Development (GED) certificate];

- **Parenting education:** high-quality instructional programs to empower parents to support the educational growth of their children; and

- **Early childhood education:** developmentally appropriate educational services for children designed to prepare children for success in regular school programs.

Each family is required to participate in all three core services.

The Even Start program has the potential to benefit families in several domains. As illustrated in Exhibit 1.1, a conceptual model of how Even Start works, the desired outcomes for parents include positive effects in three areas linked to the Even Start objectives or goals: literacy behaviors (e.g., shared literacy events with children, increased reading and writing activities in the home), parenting behavior and skills (e.g., positive parent-child relationships, positive expectations for child), and educational skills (e.g., improved reading and English language ability, higher educational attainment). In addition, goals for parents participating in Even Start might include growth in personal skills (e.g., increased self-efficacy) and community involvement (e.g., increased involvement in schools).

Ideally, Even Start will have a positive impact on children's school readiness and school achievement. School readiness includes age-appropriate cognitive, language, and social skills. Once children enter school, outcomes might include satisfactory school performance and improved school attendance, as well as a lower incidence of special education, remedial placement, and retention in grade.

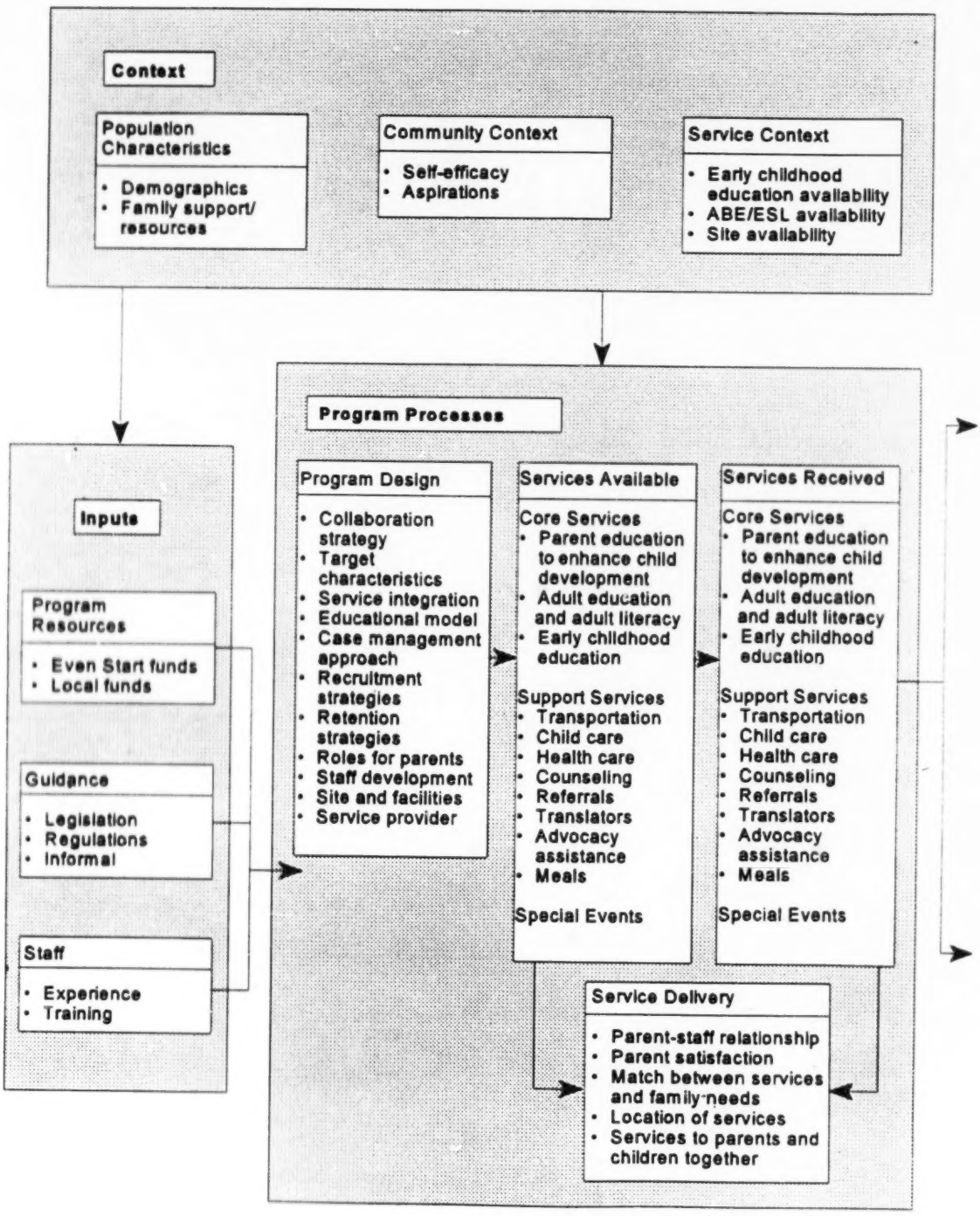
Summary of the First National Even Start Evaluation Design

The Even Start legislation requires an independent evaluation of the projects funded under Even Start. In January 1990, the U.S. Department of Education (ED) awarded a contract to Abt Associates Inc., with a subcontract to RMC Research Corporation, for the first evaluation of the Even Start program. This

[1] In April 1996, the Even Start statute was amended to require high-quality, intensive instructional programs in the areas of adult and parenting education. This new requirement became effective for projects in program year 1996-97.

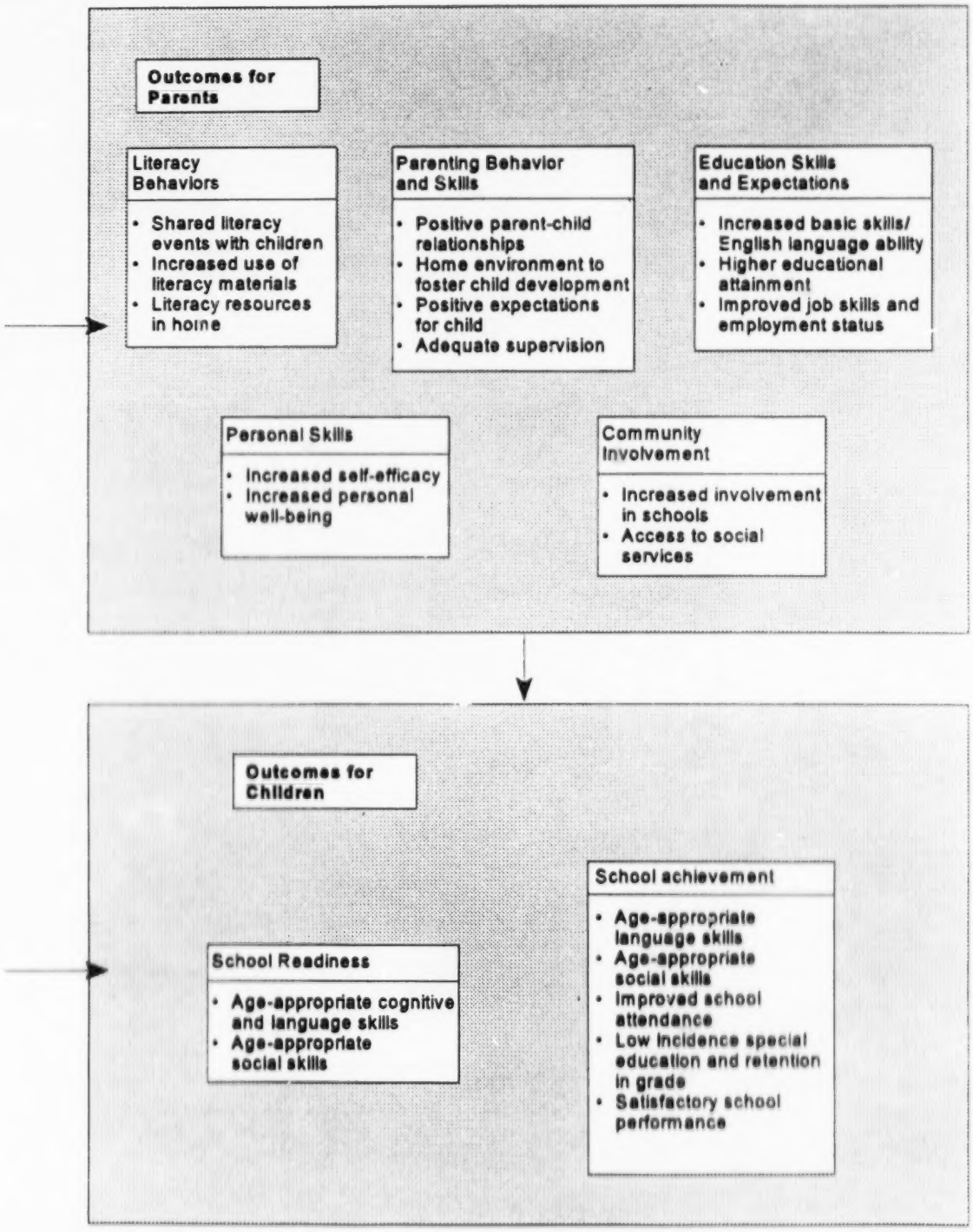
Exhibit 1.1

Even Start Conceptual Model



BEST COPY AVAILABLE

Exhibit 1.1
(continued)


Outcomes for Parents
Literacy Behaviors
• Shared literacy events with children
• Increased use of literacy materials
• Literacy resources in home
Parenting Behavior and Skills
• Positive parent-child relationships
• Home environment to foster child development
• Positive expectations for child
• Adequate supervision
Education Skills and Expectations
• Increased basic skills/ English language ability
• Higher educational attainment
• Improved job skills and employment status
Personal Skills
• Increased self-efficacy
• Increased personal well-being
Community Involvement
• Increased involvement in schools
• Access to social services
Outcomes for Children
School Readiness
• Age-appropriate cognitive and language skills
• Age-appropriate social skills
School achievement
• Age-appropriate language skills
• Age-appropriate social skills
• Improved school attendance
• Low incidence special education and retention in grade
• Satisfactory school performance

BEST COPY AVAILABLE

evaluation, which ran from 1990 through April of 1994, resulted in the design and implementation of the following four-part evaluation, with annual reports to ED and a final report to Congress:

- What are the characteristics of Even Start participants? Who participates in the program?
- How are Even Start projects implemented and what services do they provide?
- What Even Start services are received by participating families? What do families receive by participating?
- What are the effects of Even Start projects on participants? (What difference does Even Start make in the lives of participants?)

To address these questions, the national Even Start evaluation included four components. Two of the four components were carried out by the evaluation contractors:

- the National Evaluation Information System (NEIS) to collect information about program services and participants from all Even Start projects; and
- an In-Depth Study to collect more detailed information about program services and program outcomes in ten projects.

In addition to these components, projects were required to conduct local evaluation activities and were encouraged to apply for entry into the National Diffusion Network.

The In-Depth Study (IDS) was conducted in ten projects from the first cohort of 73 Even Start grantees. The projects were selected on the basis of geographic location, level of program implementation and willingness to participate.

In five of the IDS projects, families were randomly assigned to Even Start or a control group. In each project, approximately 40 families who were eligible for Even Start and had a child three or four years old were recruited by local project staff and signed consent forms to indicate their willingness to participate in the study. Almost 200 families were included across the five sites--100 Even Start families and 94 control families. Data were collected on a target adult and a target child in each of these families at three time periods: (1) upon entry to Even Start, in the fall/winter of 1991; (2) approximately nine months after entry, in the spring/summer of 1992; and (3) about 18 months after entry, in the spring of 1993. The analyses of data from the NEIS and the IDS were conducted in the summer of 1993 and the final report was submitted by Abt Associates and RMC Research Corporation to ED in the spring of 1995.[2]

Highlights of Findings from the In-Depth Study of the First Evaluation

The main objectives of the IDS were focused on initial and short-term assessment of Even Start's effectiveness on child and parent literacy. Some of the key findings from the Final Report are listed below:

[2]The report, *National Evaluation of the Even Start Family Literacy Program: Final Report, 1995*, is available from the U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service.

- More Even Start adults obtained GED certificates than did the comparison adults; 22 percent of Even Start participants obtained GED certificates during their Even Start tenure, compared to approximately 6 percent of adult participants in the control families.

- Overall, Even Start children made learning gains as a result of their own and their parents' participation in Even Start services. Further, Even Start children achieved greater gains on measures of school readiness relative to control group children at the first follow-up, although the control group children caught up with Even Start children once they entered kindergarten.

- The amount of early childhood education actually received by children made a significant difference in learning gains attained on the vocabulary test. Further, the amount of parenting education received by parents resulted in additional gains by their children on the vocabulary test, beyond those gains due to the early childhood education alone.

Research Questions for the Follow-Up Study

The In-Depth Study of the Even Start National Evaluation focused on short-term impact. The purpose of this follow-up study is to gather information about how Even Start children are performing now that they are in the primary grades, to begin to investigate medium-term effects of the program on participating children. The primary research question for this follow-up study is: What is the impact of the Even Start program on participating children in terms of their academic performance in the early elementary grades?

A secondary research question is: What are the characteristics of the schools in which the Even Start and comparison children are enrolled? Although our primary focus remains on outcomes for children, we must also take their current school and district contexts into consideration.

Design of the Follow-Up Study

In order to address these research questions, we have collected follow-up data on a subset of children from the In-Depth Study of the first national Even Start evaluation (See Exhibit 1.2). Of the 179 children included in the random assignment component of the IDS, we have been able to find, obtain parental consent and collect information on 128 children (or 72 percent) for the Follow-Up Study (63 families and 65 Even Start families).

In order to collect new data on the children in the Follow-Up Study, there were three primary tasks. First, senior staff from Abt Associates visited each of the five projects from the original In-Depth Study to discuss the Follow-Up Study with project and school district staff. We then hired staff from the Even Start project or school district to contact families, determine which school district and building each child attended, and obtain parental permission for accessing the student records. Once the children and their respective schools were identified, Abt staff visited the schools to collect data from the children's school records; staff also learned about school-level policies such as promotion/retention, availability of Title I services, English-as-a-second-language instruction, and other services. These data were collected in the spring of 1996, approximately 54 months after the children first entered Even Start. We collected data from children's 1994-1995 school year, when most children were in first and second grades, to ensure that we had complete data on attendance, grades, and achievement tests. By comparing the academic performance of participating

children in Even Start with the academic performance of children who have not participated in the Even Start program, we hope to estimate the medium-term impact of the Even Start program.

Exhibit 1.2

Number of Even Start and Comparison Children, by Project Site

Project Site	In-Depth Study (n=179)			Follow-Up Study (n=128)		
	Number of Even Start Children	Number of Comparison Children	Total	Number of Even Start Children	Number of Comparison Children	Total
Birmingham, AL	21	18	39	15	14	29
Phoenix, AZ	18	18	36	10	12	22
Golden, CO	18	14	32	10	12	22
Albuquerque, NM	18	19	37	14	14	28
Reading, PA	19	16	35	16	11	27
Total	94	85	179	65	63	128

Exhibit reads: Of the 19 Even Start and 16 comparison children from Reading who participated in the In-Depth Study, there are 16 Even Start and 11 comparison children participating in the Follow-up Study.

In our analyses we use selected pieces of data collected during the In-Depth Study. This includes basic demographic data obtained during parent interviews (e.g., age or birthdate, gender, and race/ethnicity), outcome measures (e.g., the child and adult tests), and projects' contact logs (e.g., number of hours in each core service area for participating families).

The remainder of this report is organized as follows: Chapter Two explores the hypothesized effects of Even Start as an example of an intergenerational family literacy program based on what we know from previous research and evaluation studies. Chapter Three describes what we learned both about the effects of Even Start on children and what we learned about the study participants' current school environments. The Methodological Appendices describe how we collected follow-up data in the field, and outlines the measures and analytic approaches we used.

Chapter Two

HYPOTHESIZED EFFECTS OF EVEN START

What would we anticipate as evidence of the medium-term impact of Even Start for children? We would hope that children's participation in Even Start would have a positive impact on their academic achievement in the early elementary grades; we would also hope to find that parents' participation in parenting education would positively affect their children's school performance. We set out to test these hypotheses by asking the following questions:

- What do we already know about the longer term effects of early childhood education from studies of early childhood education programs, including those with parenting education components, and two-generation programs?
- Based on the evidence from those studies, what might we expect to discover of Even Start children's early school performance?

The Even Start program approach suggests that positive outcomes on children's school readiness and school achievement are expected to occur as a result of program participation. We know from the first National Even Start Evaluation that Even Start's multi-strategic approach of providing adult education skills, early childhood education, and parenting education affects children's *school readiness in the short term*, specifically on measures of language development and vocabulary.[3] In terms of children's *school achievement*, the Even Start approach suggests that parental gains in literacy levels may have ongoing potential for affecting their children's literacy development and school achievement as measured by the following: age-appropriate language and social skills; improved school attendance; low incidence of special education, remedial placement, and grade retention; and satisfactory school performance.

We conducted a comprehensive research review of early childhood education programs to learn about the longer term impact of such programs on children's academic achievement. We focused our review studies on programs with similar goals and populations as Even Start. The number of studies reviewed was limited to include those studies that offered evidence of the effects of preschool educational programs on at-risk children's school success in the elementary grades.

In the last decade there has been a transition in the focus of programs that targeted primarily the child or primarily the parent. Early studies of preschool programs in the 1960s and early 1970s, whether researcher-designed or large-scale public programs, concentrated on measuring children's cognitive performance using traditional IQ tests (Goodson and Hess, 1978; Lazar and Darlington, 1982; Madden et al., 1976). In the 1970s and 1980s, other programs were designed to help children more broadly by focusing efforts on improving parenting skills and thus children's development as a result. Lacking sufficient evidence that either preschool education or parent education alone significantly and positively affects children's school success, researchers and program designers began targeting programs simultaneously on both generations, labelled two-generation programs (St. Pierre et al., 1995b). The transition to two-generation programs has been paralleled by a gradual shift away from evaluation studies that use cognitive measures alone to studies that use broader social,

[3] As measured by post-tests on the Peabody Picture Vocabulary Test-Revised (PPVT-R) 9 months after program entry (St. Pierre et al., 1995a).

emotional, and behavioral measures to examine children's early development and later school success (Lally et al., 1988; St.Pierre et al., 1994, 1995b; Weikart, Bond, & McNeil, 1978).

Evidence of Short-Term Effects of Preschool Education Programs

There is a substantial body of research demonstrating that early childhood educational interventions for at-risk preschoolers have positive short-term effects on children's cognitive abilities and school readiness (Goodson and Hess, 1978; Madden et al., 1976; Seitz et al., 1985; Slaughter, 1983). Although immediate positive effects were revealed by those early studies, many follow-up studies of Head Start (St. Pierre et al., 1994) and other early intervention programs have offered strong evidence that initial positive IQ gains are not maintained by participant children in later years (Casto and Mastropieri, 1986; McKey et al., 1985; Seitz et al., 1985). Despite strong evidence of the IQ fade-out effect, later research illustrated that preschool education benefits children in other ways, specifically by helping them acquire a range of skills necessary to succeed in the early elementary classroom (Weikart, 1987).

Evidence of Long-Term Effects of Preschool Education Programs

There is a growing body of research that suggests there are *long-term* positive consequences of high-quality, intensive preschool education on children's *later* school success. Those preschool program studies that were rigorously designed and that included randomly assigned control groups offer greater reliability of results (Barnett, 1992). For example, the Ypsilanti Perry Preschool Project found significant positive effects on participant children's IQ score gains, school achievement, and later success (Weikart et al., 1978). The Perry follow-up study through grade four reported that experimental group children were judged significantly more successful than control group children on meeting the demands of school, based on measures of performance on achievement tests, teacher ratings of their socio-emotional adjustment, and actual success in school in terms of grade and class placement (Weikart et al., 1978). Additionally, differences in academic achievement between the experimental and control groups actually increased over time, particularly on the California Achievement Test (CAT). Significant differences on CAT subtests gradually increased from first through fourth grades. Of particular relevance to the Even Start Follow-Up Study, the Perry follow-up found no significant differences between groups in actual school success during first or second grades, and few children from either group were retained or placed in special education.

The Abecedarian Project follow-up study also found positive long-term effects. Follow-up data indicated positive differences in intellectual development and ability for program children at several ages (Campbell and Ramey, 1994). In addition, there were significant differences in grade retention but not in special education placement or other compensatory services during the children's first three school years. (Children in this project received intensive full-day, center-based day care and preschool five days a week for five years, which is a significantly longer and more intensive intervention than Even Start children received.)

Other studies also report long-term effects. A comprehensive examination of over 200 Head Start studies found that children who participated in Head Start outperformed nonparticipants on long-term school measures such as grade retention and special education placement rates (McKey et al., 1985). Also, the Consortium for Longitudinal Studies found that early education programs for children from low-income families had long-lasting effects in four areas: school competence, developed abilities (chiefly performance on cognitive measures), children's attitudes and values, and impact on the family (Lazar and Darlington, 1982). The consortium reported that children who participated in early childhood programs were *significantly* less likely than control group children to be assigned to special education classes and *somewhat* less likely to be retained

in grade (1982).[4] Other studies investigating long-term effects also found statistically significant effects on program participants' rates of grade retention or special education placement (Campbell and Ramey, 1994; Seitz et al., 1985).

According to Barnett, a conservative interpretation of all studies' reporting long-term effects would argue that only the Perry Preschool Project and the Abecedarian follow-up results offer significant evidence of positive impact on children's school success (1995). Simultaneously, he notes that such an interpretation would ignore important information such as context and other populations served derived from other, less rigorously designed studies.

Two-Generation Program Effects

Two-generation programs are relatively new multi-strategic programs that offer high-quality early childhood education to children from low-income families as well as literacy, education, parenting and/or job training skills to their parents to enhance their opportunities for economic self-sufficiency. The Parent and Child Development Centers (PCDCs) of the early to mid-1970s were the forerunners to the current two-generation programs. As with other early interventions, the PCDCs demonstrated positive short-term effects on children, but only the second follow-up evaluation of the Houston PCDC found teachers' ratings of better school behaviors to be statistically significant (Johnson, 1989). Although the PCDCs and current intergenerational interventions have documented some short-term program effects, there is no evidence illustrating that positive program impacts on parents' parenting skills and basic education will translate eventually to better school success for their children.

A number of studies of two-generation programs, including Avance Family Support and Education Program, Child Family Resource Program, and New Chance, as well as studies of the Comprehensive Child Development Program and Head Start Family Service Centers, do document some program effects (see review in St. Pierre et al., 1995b). The studies cited in this review used a randomized experimental design, and all documented short-term positive effects on parenting measures, parent involvement, and adult educational enrollment or GED attainment. However, none of these programs has had a follow-up study to determine whether short-term positive impacts on parenting measures and educational enrollment translate to children's school success as a result of the family support/literacy program. The current gap in knowledge about the long-term effects of two-generation intervention programs for at-risk populations reflects the relative youth of the current, large-scale two-generation program models, and the lack of sufficient longitudinal data with which to evaluate long-term effects.[5]

---

[4] However, the findings from the Consortium's follow-up study are derived from pooled data on small individual samples of children from programs that began in the 1960s, which suggests that its findings should be treated conservatively when attempting to generalize to this Follow-Up study.

[5] Evaluations of the Kenan Trust Family Literacy program and Kentucky's Parent and Child Education (PACE) program have included some follow-up studies, but have not used a randomized design. One study of preschool participants who were at risk of failure when they enrolled in the Kenan Trust Family Literacy program showed that those students were ranked higher by teachers and parents than their classmates on a set of behaviors seen as important for success in school (see Seamon et al., 1991). Positive effects on child behaviors were found for PACE participants based on teacher ratings of classroom behaviors, motivation, and attendance (see Devlin, 1993), but because the evaluation was non-experimental its findings are not conclusive.

Summary

Our review suggests that early childhood educational programs that target at-risk children can and do have longer-term positive effects. Evidence from the most rigorously designed studies indicates that participating children demonstrate improved academic performance, greater social competence, and higher motivation in school than do comparison children. Less conclusive evidence from non-experimental studies also suggests positive effects on children's school-related behaviors and performance. More recent two-generation studies are as yet inconclusive, largely because such programs have not been in existence long enough to have collected longitudinal data. On the basis of our review, then, we might expect to find effects of Even Start on children's school-related behaviors and performance.

There are two features of the available evidence that are especially relevant to hypothesized effects of Even Start: one, the length of time between the program and subsequent measurement points, and two, sources of data. First, many of these studies collected data on children several years after exposure to the program, when children had completed upper elementary or even later school grades. Second, those studies that found positive effects often collected information from a variety of sources, including teachers, parents, and the children themselves. It is important to keep these features in mind as we describe what we learned in this Follow-Up Study.

Chapter Three
FINDINGS FROM THE FOLLOW-UP STUDY

This chapter describes what we learned about the 128 Even Start and comparison children we found as well as what we learned about their school and district contexts. We compared Even Start and comparison children on such measures as attendance, grades, and participation in special programs (specific child-level data and school-level data elements are described in Appendix B). Even Start children are significantly less likely to come to school late than are comparison children; otherwise we found little evidence of any medium-term effects of Even Start on children's academic records, attendance, or participation in special programs.

School Environments

Initially, the participants in the IDS were located in one of five school districts. At the time of the Follow-Up Study, the 65 Even Start and 63 comparison children were enrolled in 72 schools situated in 21 school districts during the 1994-1995 school year. To help readers better interpret the student-level outcomes presented later in this chapter, we describe what we found for children in light of variation in the schoolwide policies of these schools. All of the information presented in this chapter was drawn from children's school records for 1994-95 and from the school-level data about programs and policies in effect for the same school year.

The schools attended by Follow-Up Study participants are relatively homogenous with respect to several features: the length of the school year, the percentage of poverty, and the use of heterogenous grouping of students in primary grades. The majority of these schools are in session for between 176 and 180 days, although nearly 90 percent of schools extend their regular school year calendars by an average of 19 days. Of those schools with summer school programs, 63 percent were voluntary for all students and 37 percent were mandatory for some students and voluntary for others.

Over 75 percent of the schools have poverty levels close to or above 60 percent, and the average across the schools is 70 percent of students eligible for free or reduced-price lunch. Nearly all of the schools had some sort of heterogeneous grouping according to the student ability level (99 percent). Of the schools that grouped by ability, 18 percent did so only for some classes.

Student Demographics

The majority of the children in the Follow-Up Study were in the first and second grades during the 1994-95 school year. When these children entered the In-Depth Study several years ago, they were between 3 and 5 years old, and at the time of the Follow-Up Study (e.g., the beginning of the 1994-95 school year, approximately two-and-a-half years after the last post-test in the IDS), the majority of children were between 6 and 8 years old.

Participation in Special Programs/Services

Eighty-six percent of the schools offered Title I services in reading/language arts or math to at least one of the grade levels, and approximately one-third of schools had a schoolwide Title I project that allowed them to serve all students in the school (Exhibit 3.1). Most schools (about 90 percent) used test scores and/or teacher

BEST COPY AVAILABLE

Exhibit 3.1
Percent of Schools with Title I Services
1994-1995 School Year
(n=72) [6]

Grade Level	Percent with Reading/Language Arts	Percent with Math
Kindergarten	69	38
First	85	42
Second	88	43
Third	86	53

Exhibit reads: Sixty-nine percent of the schools reported that they offer Title I services in reading/language arts to kindergarten students.

recommendations to place children in Title I, and about one-third of the schools used grades or other criteria. Most schools reported using multiple criteria. Twenty-nine percent of schools offered some other compensatory education program in reading/language arts or math to at least one of the grade levels.

Of the children enrolled in schools that offered Title I services, 53 percent of Even Start and 53 percent of comparison children participated in Title I reading (Exhibit 3.2). Nearly seventy percent of the schools reported that they offer Title I reading services in kindergarten, yet only about one-third offer services in math. This has potential ramifications for Even Start children who move into public school; if there are limited Title I services available, children may lose a potentially valuable bridge between Even Start and school during a critical transition year from early childhood education into the public school system. However, there were no significant differences between Even Start and comparison children for either Title I reading or Title I math service participation.

There were no differences in the percentage of students identified as special needs (9 percent of Even Start and 10 percent of comparison children); this proportion parallels the proportion found in the In Depth Study.

Schools varied in terms of offering other special programs, such as transitional classrooms between grade levels, extended year programs, and bilingual instruction. Exhibit 3.3 indicates participation rates for Even Start and comparison children (the number of students eligible to receive a given service is indicated in parentheses below the percent). There are no significant differences between Even Start and comparison children in participation in other special programs.

[6]Two of the schools in the sample did not have a third grade level, so the description of policies at the third grade level was calculated with a sample size of 70 schools.

Exhibit 3.2
Children Participating in Title I Services
1994-1995 School Year

Receipt of Title I Services	Percent Even Start		Percent Comparison	
	Reading (n=55)	Math [7] (n=21)	Reading (n=53)	Math (n=22)
Yes, received Title I	53	62	53	59
No, did not receive Title I services	47	38	47	41

Exhibit reads: Forty-seven percent of both Even Start and comparison children did not receive Title -- I services in reading.

Exhibit 3.3
Children Participating in Available Special Programs
1994-1995 School Year

Program	Percent Even Start (number eligible)	Percent Comparison (number eligible)
Bilingual education	59 (41)	53 (36)
Compensatory education (reading)	22 (9)	17 (12)
ESL services	16 (43)	26 (39)
Summer school	8 (59)	8 (52)
Transitional classroom	0 (16)	18 (11)

Exhibit reads: Twenty-two percent of Even Start children (of the 9 children eligible) received compensatory education services in reading (other than Title I).

[7] Fewer schools offered Title I services in math; 25 schools (of the 62 that offered any Title I services) offered instruction in math.

21 BEST COPY AVAILABLE

Student Attendance and Performance

The attendance rate for each child was calculated as the total number of days the child was present in school for the year divided by the total number of days school was in session. The average attendance rate (94 percent) did not differ for children in the Even Start and comparison group (see Exhibit 3.4). There were also no significant differences in the number of absences between the two groups. The average tardiness rate was similarly calculated (the total number of tardy arrivals reported for each child divided by the total number of days the child was present in school for the year). The average tardy rate for the Even Start group was 1 percent, and the tardy rate for the comparison group was 4 percent; this difference was statistically significant at the 0.05 level.

Schools reported that in the early grades, typically kindergarten, first and second grades, they generally characterized students' progress descriptively, using language such as excellent, satisfactory plus, mostly satisfactory, adequate performance, and improving or emerging in skills. Seventy-five percent of the schools reported using letter grades in third and higher grades, and the use of descriptive grades began to trail off in the middle elementary grades. There are no differences between the grades of Even Start and comparison children for reading, language arts, or math at any grade level (Exhibit 3.5). In fact, both Even Start and comparison students seemed to perform at average or above average levels. There were also no differences in grades when controlling for number of hours in preschool education, number of parents' hours in adult education or parenting education, or children's scores on the Preschool Inventory (PSI) or the Peabody Picture Vocabulary Test (Revised) (PPVT-R), all of which were collected in the In Depth Study.

Achievement Tests

Very few schools reported that they used achievement tests to assess young children's progress. Only two schools offered achievement tests at the kindergarten level, for example. A larger proportion of schools (72 percent) indicated that they used achievement tests for second grade. However, schools vary considerably not only in the types and levels of the tests used, but also in the purposes for testing. Some schools administer tests to assess performance of all students at a given grade level (typically grades 3 or 4), while others use achievement tests only in order to identify students in need of instructional services such as Title I or other compensatory education. Among the different types of achievement tests used are the Brigance, Metropolitan Achievement Tests (MAT), California Achievement Tests (CAT), Iowa Tests of Basic Skills (ITBS), Stanford Achievement Test (SAT), and the Otis-Lennon Scholastic Aptitude Test (OLSAT). While some schools reported that they conducted gradewide assessments of primary school students, many used instruments for diagnostic or school readiness purposes rather than for assessing students' achievement. In fact, 20 schools reported that they use standardized tests primarily for diagnostic purposes. Only 34 students (including first, second, and third graders) were tested on any reading achievement test, and the reasons for testing those students included both grade-level and diagnostic purposes. There were no differences between Even Start and comparison children who were assessed with the same test.

Summary

In this study, the medium-term effects of Even Start on children (in grades 1 and 2, about three years after participating in Even Start) were assessed using the following measures: school grades; achievement test data; school attendance; special education placement; Title I placement; participation in transitional classrooms, summer school programs and bilingual education; and grade retention history. In addition to these outcome measures descriptive information was obtained on the Follow-Up Study children such as special needs status. These measures were drawn from the school records of children in both the Even Start and comparison groups

Exhibit 3.4
Attendance and Tardiness Rates

1994-95 Average Rates of	Even Start (n = 65)	Comparison (n = 63)
Attendance	94%	94%
Tardiness	1%	4%*

*p < .05

Exhibit reads: The average attendance rate for both Even Start and comparison children was 94 percent.

Exhibit 3.5
Average Grades in Reading, English, and Math

Content Area	Even Start (n = 65)	Comparison (n = 63)
Reading	1.8	1.8
English	1.8	1.8
Math	1.6	1.7

Exhibit reads: The average reading grades for both Even Start and comparison children were 1.8 on a 3-point scale.

Note: Grades were converted to a 3 point scale (corresponding to (1) very good, (2) satisfactory, and (3) needs improvement) from descriptive, narrative, and other scales solely for the purpose of comparing academic performance.

during the 1994-95 school year, which was approximately 54 months after pretest. A total of 128 student records were assessed which includes 65 Even Start children and 63 comparison children. This sample size would allow us to detect a medium-sized effect of treatment.

At this time, and with the available measures, we found essentially no differences between Even Start and comparison children, with the exception of a higher tardiness rate for comparison children. Given what we know from the research literature about the long-term effects of early childhood education programs, however, it is not surprising that we found virtually no medium-term effects of Even Start. Studies that found long-term effects had several features the current Follow-Up Study did not, including: a) a longer interval between the treatment and subsequent follow-up, typically when the children had completed fifth (or higher) grade; b) the use of comparable grade data and/or achievement test scores; and c) the capacity to collect new data from children as well as from teachers and/or parents. Many studies that find long-term effects looked at measures of children's classroom behavior (as assessed by teachers) and social competence (as assessed by parents and teachers) as well as out-of-school behaviors. While it is certainly possible that the Even Start program does not have medium- or long-term effects (as measured in this study), it is also possible that we might be able to detect meaningful differences were we able to meet the conditions described above

REFERENCES

Barnett, W.S. (Winter 1995). Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. In *The Future of Children: Long-Term Effects of Early Childhood Programs*. Los Altos, CA: David and Lucille Packard Foundation, The Center for the Future of Children. 5, 3, 25-50.

Barnett, W.S. (1992). Benefits of Compensatory Preschool Education. *The Journal of Human Resources*. 27(2), 279-312.

Campbell, F.A., and Ramey, C.T. (1994). Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children From Low-Income Families. *Child Development*, 65(2), 684-698.

Casto, G., and Mastropieri, M.O. (1986). The Efficacy of Early Intervention Programs: A Meta-analysis. *Exceptional Children*. 52(5), 417-424.

Devlin, K.M. (1993). Innovations: Kentucky's Parent and Child Education (PACE) Program. Cabinet for Workforce Development, Department of Adult and Technical Education. Frankfort, KY: Author.

Goodson, B.D. and Hess, R.D. (1978). The Effects of Parent Training Programs on Child Performance and Behavior. In B. Brown (ed.), *Found: Long-term Gains From Early Intervention*. AAAS Selected Symposia Series, No. 8. Boulder, CO: Westview Press.

Johnson, D. (1989). The Houston Parent Child Development Center Project: Disseminating a Viable Program for Enhancing At-Risk Families. In R.P. Lorian (ed.), *Protecting the Children: Strategies for Optimizing Emotional and Behavioral Development*. Prevention In Human Services, 7, New York, NY: Haworth Press.

Johnson, D., and Walker, T. (1991). *Final Report of an Evaluation of the Avance Parent Education and Family Support Program*. Report submitted to the Carnegie Corporation. San Antonio, TX: Avance.

Lally, J.R., Mangionne, P.L., and Honig, A.S. (1988). The Syracuse University Family Development Research Program: Long-Range Impact on an Early Intervention with Low-Income Children and Their Families. In Powell, D. (ed.) *Parent Education as Early Childhood Intervention*. Norwood, NJ: Ablex.

Lazar, I., and Darlington, R. (1982). Lasting Effects of Early Education: A Report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, 47(2-3), Serial No. 195.

Madden, J., Levenstein, P., and Levenstein, S. (1976). Longitudinal IQ Outcomes of the Mother-Child Home Program, 1967-1973. *Child Development*, 47, 1015-1025.

Marcon, R.A. (March 1993). *Parental Involvement and Early School Success: Following the 'Class of 2000' at Year Five*. Paper presented at the Biennial Meeting of the Society for Research in Child Development (60th, New Orleans, LA).

REFERENCES
(Continued)

McKey, R.H., Condelli, L., Ganson, H., Barrett, B., McConkey, C., and Plantz, M. (1985). *The Impact of Head Start on Children, Families and Communities: Final Report of the Head Start Evaluation, Synthesis, and Utilization Project*. Washington, DC: CSR, Incorporated.

Quint, J.C., Polit, D.F., Bos, H., and Cave, G. (1994). *New Chance: Interim Findings on a Comprehensive Program for Disadvantaged Young Mothers and Their Children*. New York, NY: Manpower Demonstration Research Corporation, 1994.

Seamon, D., Popp, B., and Darling, S. (1991). *Follow-up Study of the Impact of the Kenan Trust Model for Family Literacy*. Louisville, KY: National Center for Family Literacy.

Seitz, V., Rosenbaum, L.K., and Apfel, N.H. (1985). Effects of Family Support Intervention: A Ten-Year Follow Up. *Child Development*, 56, 376-391.

Slaughter, D.T. (1983). *Early Intervention and its Effects on Maternal and Child Development*. Monographs of the Society for Research in Child Development, 48(4), Serial No. 202.

St. Pierre, R.G., Swartz, J., Gamse, B., Murray, S., Deck, D, and Nickel, P. (January 1995). *National Evaluation of the Even Start Family Literacy Program: Final Report*, prepared for the U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service. Cambridge, MA: Abt Associates, Inc.

St. Pierre, R.G., Layzer, J., and Barnes, H.V. (Winter 1995). Two-Generation Programs: Design, Cost, and Short-Term Effectiveness. In *The Future of Children: Long-Term Effects of Early Childhood Programs*. Los Altos, CA: David and Lucille Packard Foundation, The Center for the Future of Children. 5, 3, 76-93.

St. Pierre, R., Goodson, B., Layzer, J., and Bernstein, L. (May 1994). *National Impact Evaluation of the Comprehensive Child Development Program: Report to Congress*. Cambridge, MA: Abt Associates, Inc.

Swartz, J.P., Smith, C., Berghauer, G., et al. (1994). *Evaluation of the Head Start Family Service Center Demonstration Projects: First Year Evaluation Results*. Cambridge, MA: Abt Associates, Inc.

Travers, J., Nauta, M., and Irwin, N. (1982). *The Effects of a Social Program: Final Report of the Child and Family Resource Program's Infant Toddler Component*. Cambridge, MA: Abt Associates, Inc.

Weikart, D.P., Bond, J.T., and McNeil, J.T. (1978). *The Ypsilanti Perry Preschool Project: Preschool years and Longitudinal Results Through Fourth Grade*. (Monograph 3), Ypsilanti, MI: High/Scope Educational Research Foundation.

Weikart, D.P. (1987). Curriculum Quality in Early Education. In S.L. Kagan and E.F. Zigler (eds.), *Early Schooling: The National Debate*. New Haven, CT: Yale University Press.

Appendix A
FINDING FAMILIES

As described earlier, the In-Depth Study had three data collection points: testing at entry to Even Start, and retesting again at nine and eighteen months after entry. To be eligible for the Follow-Up Study, families had to have been pretested. If the target child had no baseline data, s/he was considered ineligible for the current study. Of all the IDS families in the five sites with random assignment, 179 Even Start and comparison children had at least one session of testing recorded, which made them eligible for the Follow-Up Study. Once this subset of the original comparison and Even Start groups had been identified, we were able to begin work on locating the families.

Preparation of Family Fact Sheets

At each of the data collection points in the IDS, field staff updated each family's address and telephone data, and also obtained the names, telephone numbers, and addresses for two additional people who would most likely know where IDS participants could be reached in the future. For families who had remained in the IDS throughout the data collection, address and telephone information had been updated as recently as 1993, while the address information was less current for others. We used our 1993 addresses as the jumping off point for our efforts to locate all the families.

Initial Contact with Sites/Local Recruiters

We visited each of the five IDS projects to discuss the follow-up effort with Even Start staff and school district personnel, to explain the purpose of the Follow-Up Study and to request their assistance in hiring an on-site local recruiter to locate families. In all five sites, project directors suggested that someone familiar with the community, language, and environment of the sites would be more effective in making contact and gaining consent from families. Furthermore, given the outdated addresses and phone numbers, we assumed that word-of-mouth leads from people in the community might be a valuable resource for tracking down hard-to-reach families. All five Even Start Project Directors chose either part-time staff or former Even Start parents to hold the temporary position of on-site recruiter.

In late 1995 and early 1996, Abt senior staff conducted one-day site visits to the five IDS sites in order to meet with district and Even Start staff interested in the Follow-Up Study, to present an outline of the study goals and timeline, an overview of the data collection instruments we had drafted, and a copy of the parent consent forms for their comments and approval. We also solicited input from district/school personnel about our draft record extraction forms.

District staff were unanimously willing to facilitate the study efforts and offered their cooperation readily to Abt staff, particularly in identifying which of the original students remained in their school district. They also informed Abt staff of the availability of the data sought, the location where the information would be housed (e.g., centrally at the district office or at the individual schools), and the format in which it was maintained (e.g., on diskettes or tape drives versus hard copy records). In addition to meeting with the Even Start and district staff, Abt staff interviewed and selected the candidates for local recruiters who had been recommended by Even Start project directors. Abt staff conducted separate training meetings with the recruiters and the Even Start project staff; these training sessions covered training materials and resources, reviewing the

BEST COPY AVAILABLE

addresses we had for families, consent forms to be used in the field, and forms for updating addresses and consent form receipts.

From information gathered at these meetings, Abt staff were able to refine the data collection forms and develop plans for on-site data collection with the on-site staff in each of the five districts. Additionally, we learned that a number of families were either continuing participants in Even Start or had remained in contact with Even Start project staff. Interestingly, we also learned that two or three families *had* at one time been comparison families and had subsequently enrolled in Even Start[8].

Locating Families and Obtaining Consent

Once Abt staff had completed visits to all five projects, the local recruiter(s) in that site began to locate families. Among the questions the recruiters asked parents was which school and district the target child had attended during the 1994-95 school year and which grade the child had been in (during the 1994-95 school year). The recruiters also asked parents whether they would permit Abt staff to review their children's school records.

When the recruiters found families and parents had given verbal consent, they arranged meetings either at the families' homes or in other convenient locations (such as local schools or Even Start offices). Each family received a description of the study (in English or Spanish, as appropriate) and the recruiters asked parents to sign a consent form granting us permission to review school records. In those instances where families had moved more than an hour's drive from the area, the recruiter sent the study description and consent form by mail (a postage-paid envelope was included in all such mailings.)

When completing the consent form, parents were asked to write in the name and birth date of their child as a verification for us that they were indeed referring to the appropriate target child in the family. Anticipating an eventual (additional) follow-up study with these same families, the recruiters also verified the accuracy of the contact information provided on the family fact sheets, updating contact data if necessary.

Data Collection

Once the recruiters had successfully contacted parents and obtained consent, we were able to plan on-site data collection in schools and districts. Because most of the families remained clustered around the same five IDS districts, we sent our staff to visit the schools and extract information from school records maintained either in individual student cumulative folders at the child's school or in electronic records maintained at the district.

Contact with the School Data Coordinators

In each of the five IDS districts, we worked with the district data coordinator to gain access to district records and to gain entry to individual schools. Across all the districts, certain data elements are routinely transferred from the school files and stored at the central district office once the school year comes to a close. The location of student data differed for each district, although most districts maintained such data as achievement scores, special education placement, and Title I designation, while schools maintained grades and attendance data. Data coordinators provided formal authorization to school staff reluctant to cooperate in the data

[8] During the In-Depth Study, some projects that agreed to random assignment consented to do so with the stipulation that after the two-year data collection period of the IDS had ended, those comparison families that were still interested in participating in Even Start could do so.

collection; they served as the reference necessary for us to establish legitimacy with the schools; and they helped clarify district policies and programs that were at times unclear from the student data.

The district data coordinators also provided us with the names and telephone numbers of the principals at the various elementary schools in the district where study children attended. We scheduled visits to districts on the basis of when schools were in session, and when the district contact person was available. In one district, both student- and school-level data were all accessible through the district computer system, thereby eliminating the need for on-site visits to the individual schools.

Visiting Schools and Districts

The second stage in our data collection consisted of visiting schools. While in each school, Abt staff obtained student data from school records as well as an overview of school-level policies from the principal or other school staff member. In those few cases where the principal wasn't available during our visits, we conducted brief telephone interviews at another time. Finally, we sent stipends to each school district that had worked with us to help collect data for the Follow-Up Study.

Appendix B
MEASUREMENT AND ANALYSIS

This section provides a comprehensive description of the methods used to collect and analyze data for the Even Start Follow-Up Study. The data collection methods and measures of the Follow-Up Study are presented first, followed by a discussion of the analytic sample and analytic methods.

Measures Used in the Follow-Up Study

Two sources of data were analyzed for this study. First, the information recently collected from the children's schools was used to assess long-term program impact. Second, we used data from the In-Depth Study as baseline measures.

Data Collected for the Follow-Up Study

To obtain new information about the children in the Follow-Up Study, we drew from two data sources: (1) children's school records, and (2) brief conversations with the principal or other key staff member in the target child's school. We collected information from the children's last school year, 1994-95, in order to obtain the most complete set of data. Standardized achievement tests, for example, are typically administered in the spring, and results are sometimes not available until summer months. Because we wanted to ensure that we had the most complete information possible, we limited our data collection to that period of time for which all test scores (and other data as well) would have been recorded.

Child-Level Information

The school records provide information on the children's academic performance during the 1994-95 school year. From the review of the individual child records, we collected information on:

- **Attendance:** The number of days the child attended school, the number of excused, unexcused and unspecified absences, number of days the child was tardy, and whether or not the child was in school for the entire year.
- **Grades:** Reported grades in reading, English/language arts, and math.
- **Standardized Achievement Test:** Achievement test scaled scores, national percentile rank and normal curve equivalents in math and reading, and type of test taken.
- **Retention:** Whether the child repeated kindergarten, first or second grade; Whether the child attended a transitional class and/or an extended year program.
- **Title I/ Other Compensatory Education:** Whether the child received Title I services or other compensatory education in reading, math or ESL.
- **Special Needs:** Whether the child has special needs, and whether the records indicated that the child had an individualized education plan.

- **Bilingual Education:** Whether the child participated in bilingual education in math or reading.

School-Level Information

Because the children in the Follow-Up Study attend schools in 21 different school districts, we collected information from key school staff, usually the principals, about different school policies in order to explain the possible influence of school procedures on the selected child outcome measures. Our questions focused on those school or district policies in place for the 1994-95 school year on the following topics:

- **Attendance:** The total number of days in the school year and whether the school distinguishes between excused/unexcused absences, or flags tardiness.

- **Grades:** The types and scales of grades (e.g., letter grades, pass/fail, descriptive information) used to assess children's progress in kindergarten, first, second and third grades.

- **Standardized Achievement Tests:** Whether the school offered standardized achievement tests in kindergarten, first grade, second or third grade and which tests are administered.

- **School Lunch Program:** The percentage of students in the school eligible for free- or reduced-price lunches, which serves as an indicator of poverty level of students in the school.

- **Retention:** The school retention policy, if any (i.e., children in primary grades are promoted in order to remain with their age-mates regardless of academic performance).

- **Title I Services:** Whether the school offered Title I services in reading, math or language arts; how children were identified for Title I; and whether the school operated a schoolwide Title I project during 1994-95.

- **Tracking:** Whether or not the school has heterogeneous grouping by ability level of children in the primary grades.

- **Transitional Classrooms:** Whether the school had any transitional classrooms that bridge traditional levels, and the criteria for placing children in transitional classrooms.

- **Bilingual Education:** Whether the school offered bilingual classrooms, how children were placed in bilingual classrooms and whether that was reflected in a student's record, and whether the school offered ESL classes.

- **Other Compensatory Education Programs:** Other compensatory education programs offered, if any (e.g., migrant education, state-specific compensatory programs).

- **Extended Year Program:** The existence of an extended year program, the selection criteria for the extended year program, whether the program was mandatory, and what grade levels were eligible for the program.

Data from the In-Depth Study

In addition to the data collected specifically for the Follow-Up Study, we use data collected on the target children and their parents at baseline of the In-Depth Study. The data collected from target adults in the In-Depth Study include an in-person interview as well as the Comprehensive Adult Student Assessment System (CASAS). The children in the In-Depth Study were administered the Preschool Inventory (PSI) and the Peabody Picture Vocabulary Test-Revised (PPVT-R).

- **Parent Interview:** An in-person interview was conducted with the target parent about questions regarding family characteristics, family resources and service utilization. Selected variables from the parent interview are used for the Follow-Up Study.

- **CASAS literacy**: The CASAS is a standardized test of adult functional literacy. The CASAS has the flexibility to measure participants involved in diverse adult education programs, spanning the range from non-readers to adults at the GED or high school level. Although the CASAS measures reading, writing, math and problem solving skills, the In-Depth Study used only the Reading Survey achievement test in order to reduce respondent burden and because Even Start was expected to have more effects in reading than in math.

- **Preschool Inventory:** The PSI was developed by Bettye Caldwell as a 64-item inventory of basic concepts important for preschool children to know before entering school (CTB/McGraw-Hill, 1970). A 32-item version, adapted for an earlier study by Abt Associates, was used in the In-Depth Study. The PSI assesses a range of school readiness skills, such as identifying shapes and colors and understanding numerical concepts.

- **Peabody Picture Vocabulary Test-Revised:** The PPVT-R measures receptive (hearing) vocabulary, and provides a quick estimate of verbal or literacy-related skills. The PPVT is an individually administered test that requires 15 to 20 minutes per child and is appropriate for children between the ages of 2 and 18 years.

Analytic Approach to Assessing Program Effects

This section describes the analytic sample used in the Follow-Up Study and our approach to assessing the medium-term effects of Even Start on the children in the study. In order to determine the medium-term effects, we focus on the difference between the academic performance of children who participated in the Even Start program and children who did not participate in the program three years after intake.

Analytic Sample

The sample for this study has three levels: Even Start projects, families associated with those Even Start projects, and the school districts that the children in the study attend. The Follow-Up Study was restricted to the five projects in the IDS that implemented a randomized design. These projects are Birmingham, Alabama; Phoenix, Arizona; Golden, Colorado; Albuquerque, New Mexico; and Reading, Pennsylvania.

Of the 179 children included in the original sample in these five projects, we were able to contact and collect information on 128 children for the Follow-Up Study (63 comparison children and 65 Even Start children). The treatment status of each child was maintained as originally assigned regardless of whether a "treatment group" family dropped out of Even Start or a "control group" family managed to enter Even Start at a later date.

The children in the Follow-Up Study come from 21 school districts and 72 different schools. Because the sample size is small, data are pooled across Even Start projects and schools.

Comparability of Follow-Up Study and In-Depth Study Samples

We were unable to locate and collect information on all of the children in the five random assignment sites in the In-Depth Study. The question then arises as to whether the Follow-Up Study sample is still representative of the sample of children and families in the In-Depth Study.

To examine this question, one sample t-tests and chi-square associations were computed on the differences between baseline characteristics of the children in the original five sites of the In-Depth Study (n=179) and the subsample of children identified in the Follow-Up Study (n=128). The baseline measures examined for these analyses included family configuration, education level of the parent or guardian, employment status of parent or guardian, primary source of family income, annual family income, gender of child, primary language of child, race/ethnicity of child, age of child, PSI, CASAS and PPVT scores. None of these differences were significant at the 0.05 level, which indicates that the two samples are statistically comparable on these characteristics.

These findings indicate that the Follow-Up Study sample and the sample from the five randomly assigned sites in the IDS are statistically equivalent in terms of the baseline characteristics listed above. However, it is still possible that the two samples vary on other unmeasured variables.

Comparability of Even Start and Comparison Groups

A second concern is whether the Even Start and comparison groups of the Follow-Up Study were significantly different from each other at baseline. To address this question, independent sample t-tests for interval or continuous variables and chi-square associations for nominal variables were conducted on these same set of baseline characteristics used in the previous analysis. There was one significant difference at the 0.05 level between the two groups. Approximately 91 percent of parents in the Even Start groups were not working at baseline compared with only 77 percent of comparison parents. On all other baseline measures, the two groups were statistically comparable. Appendix C of this report presents the distribution of the Even Start and comparison groups on these baseline variables.

Analytic Approach

The analyses for this report address the question of the medium-term effectiveness of the Even Start program on child academic performance. As noted earlier, in order to best interpret the academic measures, we

collected information from the schools attended by all children in the sample. For most outcomes, we compared the average value for Even Start children with the average value of comparison group children using independent t tests and chi-square associations.

For three of the outcome measures—grades in reading, math and English—we used standard multiple regression models to compare the Even Start and comparison groups. In the model, the grade variables were the dependent variables and the program or comparison status was the independent variable. Because we discovered significant differences between Even Start and comparison group families in one variable at baseline, we included some baseline covariates in the model to reduce variance and to increase our chances of finding significant differences. However, the small size of the sample limited the number of covariates possible.

Design Limitations

The current study differs from the earlier evaluation in several ways. First, the analytic sample of 128 used in the study does not maintain the same randomization that was used to assign families in the IDS. The Follow-Up Study sample consists only of those families/children we were able to locate and collect information from and does not incorporate the activities and behavior changes of the children in the original IDS who we were unable to locate. Further, the IDS collected data from parents as well as from children, and the current study is limited to using extant data on children. Finally, the number of school districts has increased from 5 to 21. The increased dispersion of former Even Start participants and comparison children is not in and of itself an issue, because the children were randomly assigned at the outset of the In-Depth Study. The issue is that as a result of the dispersion, the measures we use across schools are not the same.

Unlike the earlier evaluation, in which all the IDS child participants were assessed using common measures such as the Preschool Inventory and the Peabody Picture Vocabulary Test-Revised, the Follow-up Study participants were not assessed with any of the same measures. The pretest information gathered from the IDS is used to describe the sample at intake, while the new data collected for the Follow-Up Study are used to measure the academic performance at follow-up. The two types of data are not comparable and are not used in a pretest to posttest gain comparison. Because there was no pretest on the academic measures collected at follow-up, it is difficult to determine with complete confidence whether or not the Even Start and comparison groups made significantly different gains. If, for example, information on school attendance and grades had been collected at the same time point, say entry to school, we would have been able to determine if the Even Start and comparison groups were comparable at the next measurement point.

The last limitation of the study is the size of the sample. The small number of observations restricted the complexity of the analysis and limited the power to detect statistically significant differences.

Appendix C

DESCRIPTION OF EVEN START FOLLOW-UP STUDY CHILDREN AND FAMILIES AT PRETEST

This appendix describes the characteristics of the 128 families and children in the Even Start Follow-up Study at pretest during the fall/winter of 1991.[9] These data were collected for the original In-Depth Study of the first National Even Start Evaluation and include demographic characteristics such as race/ethnicity as well as scaled scores on the CASAS, PPVT and PSI.

[9]Due to missing data on some variables, the actual sample size ranges from 108 to 128.

Demographic Characteristics of In-Depth Study Families at Pretest

Exhibit C.1
Demographic Characteristics of Families at Pretest

Demographic Characteristics	Percent Even Start (n=65)	Percent Comparison (n=63)
Family Configuration		
Single Parent	40	33
Couple	46	53
Extended Family	13	14
Other	2	0
Education Level		
Parent has high school diploma, GED or	26	27
Spouse/Partner has high school diploma, GED	39	34
Employment Status		
Parent not working	91	77
Parent working	10	23
Primary Source of Income		
Government assistance	48	34
Job wages	43	57
Alimony and child support	0	2
Other	10	7
Annual Income		
Less than $5,000	51	42
5,000-9,999	22	27
10,000-14,999	14	13
15,000-19,999	6	9
20,000-24,999	3	7
25,000 or more	3	2

Exhibit reads: Forty-eight percent of Even Start families' primary source of income at pretest was government assistance.
Note: Due to rounding, column totals may not sum to 100 percent.

Literacy Level of Adults at Pretest

The literacy level of the parent or guardian of the children in the Follow-Up Study was measured by the CASAS literacy test. Of the adults tested in the Even Start and comparison groups, the mean score for both was above 225. Adults who score above a scale score of 225 can function at a high school entry level in basic reading.

Exhibit C.2
Pretest Scores on Comprehensive Adult Student Assessment System (CASAS)

	Even Start			Comparison		
	N	Mean	SD	N	Mean	SD
CASAS	47	226.43	12.8	44	228.89	16.02

Exhibit reads: Even Start adults in the Follow-Up Study averaged 226.4 points on the CASAS Reading Survey pretest.

Exhibit C.3
Children's Demographic Characteristics at Pretest

Demographic Characteristics	Percent Even Start (n=65)	Percent Comparison (n=63)
Gender		
Female	47	56
Male	53	44
Primary Language		
English	62	65
Spanish	39	35
Other	0	0
Race/Ethnicity		
African-American	29	21
White	7	11
Hispanic	62	68
Other	2	0
Age		
2-0 to 2-11	5	5
3-0 to 3-11	48	37
4-0 to 4-11	42	49
5-0 to 5-11	6	10

Exhibit reads: Forty-two percent of the Even Start children in the Follow-Up Study were between 4 years, zero months, and four years, 11 months old at the time of the pretest.
Note: Due to rounding, column tables may not sum to 100 percent.

Exhibit C.4
Pretest Scores on Preschool Inventory (PSI) and Peabody Picture Vocabulary Test-Revised (PPVT-R)

Measure	Even Start			Comparison		
	N	Mean	SD	N	Mean	SD
Preschool Inventory	54	10.74	6.13	53	11.83	5.98
PPVT-R	55	77.44	18.76	54	74.98	17.86

Exhibit reads: Even Start children in the Follow-Up Study averaged 10.7 points on the PSI pretest.